图书情报工作
第 62 卷 第 14 期 2018 年 7 月

基于主题变迁的领域发展路径智能化识别*
——以人工智能为例

■ 周源[1] 张超[2] 唐杰[3] 刘宇飞[4] 张宇韬[3]
[1] 清华大学公共管理学院 北京 100084 [2] 华中科技大学机械科学与工程学院 武汉 430074
[3] 清华大学计算机系 北京 100084 [4] 中国工程院战略咨询中心 北京 100088

摘要：［**目的/意义**］识别领域发展路径对于科技创新具有重要意义，但现有方法如专家访谈、引文分析等不能适应文献爆发性增长的现状，针对这一问题，提出一种基于主题变迁的领域发展路径识别方法。［**方法/过程**］该方法可以自动从 Aminer 平台获取数据，通过构建关键词－学者矩阵，综合使用 KMeans++ 和谱聚类算法识别出研究主题和相关学者；通过相似度计算实现不同主题之间的关联，最终获得研究领域的发展路径并进行可视化展示。［**结果/结论**］通过对人工智能领域的实证分析，结果表明该方法能够有效反映领域研究主题的变迁，有助于研究者快速定位领域的研究热点和重点，丰富领域发展路径相关的研究方法。

关键词：领域发展路径 主题变迁 KMeans++ 谱聚类 人工智能
分类号：G250.2
DOI:10.13266/j.issn.0252－3116.2018.14.008

1 引言

领域发展路径描述了领域中研究主题的产生、变迁和消亡的全过程，它可以帮助研究人员理解领域的研究历史和现状，快速识别研究的前沿热点问题[1]。学者运用德尔菲法[2]、AHP[3]等方法进行了一系列相关研究，这些方法大多是基于专家知识的定性分析，偏重于设定合适指标[4]。另有一些研究者利用主路径分析研究技术的变迁[5]，然而主路径分析仅用一个线性结构难以完整反映领域演变，会造成诸多演变细节的丢失[6]。由于德尔菲法、AHP 和主路径方法为代表的传统方法存在偏误，近些年学者们开始转向使用主题变迁进行领域发展路径的研究。主题变迁是指一个领域的研究主题的内容和热度随着时间变化的现象[7]，通过挖掘深层次的科技文本语义信息，能够有效解决传统方法存在的问题，通过可视化的方式进行展示，可以从更全面的视角感知领域发展路径。

本文提出一种基于主题变迁的领域发展轨迹智能化识别方法，数据通过 API 从 AMiner 上自动获取，无需手动下载，通过 Kmeans++ 和谱聚类联用，一方面可以控制主题的粒度，另一方面可以动态地调整主题数量。实现输入领域名称，即自动获得该领域发展路径图的智能化流程，并且从图中可以清晰看出各个时间段的研究主题和相关学者。文章的最后以人工智能领域为例进行实验以验证方法的有效性。

2 相关研究

2.1 主题变迁

主题变迁又称主题演化，通常使用数据挖掘相关方法探索主题内容和强度在时间维度上的变化以及不同主题之间的交互[8]。随着文献爆发式增长，主题变迁分析面临着数据量大和数据类型复杂的挑战，如何从海量的数据中快速准确获取领域发展脉络，是科研人员和情报人员共同关心的问题。国内外学者针对这一问题提出了很多模型，和本文相关的方法主要分为两类：一是基于聚类分析模型的方法，二是基于概率主

* 本文系国家自然科学基金“支持技术预见的多源异构大数据融合与时序文本预测方法研究”（项目编号：91646102）和国家自然科学基金“面向 2035 的中国工程科技发展路线图绘制理论与方法研究”（项目编号：L1624045）研究成果之一。
作者简介：周源（ORCID：0000－0002－9198－6586），副教授，博士生导师；张超（ORCID：0000－0001－7612－9327），硕士研究生；唐杰（ORCID：0000－0003－3487－4593），副教授，博士生导师；刘宇飞（ORCID：0000－0001－9420－8811），博士后，通讯作者，E-mail：liuyufei0418@qq.com；张宇韬（ORCID：0000－0002－5759－1230），博士研究生。
收稿日期：2017－11－21 修回日期：2018－03－14 本文起止页码：62－71 本文责任编辑：刘远颖

题模型的方法。

基于聚类分析模型的主题变迁方法主要应用于话题检测与追踪(TDT)领域和文献计量学领域。TDT将话题定义为具体时间地点发生的事件,该方法聚焦于新闻文本流的话题演化上,包括事件内容的演化和强度的演化[9]。文献计量相关方法包括引文分析、共词分析、耦合分析等,这些方法的核心思路是各种“共现分析”。崔雷等[10]使用引文共引聚类研究领域发展历史,并通过对高频词进行共现聚类分析,总结研究主题。唐果媛等[11]对国内基于共词分析法的主题演化方法进行了总结,将分析流程划分为5个步骤:确定数据源、演化阶段划分、确定分析对象、构建共词矩阵并归一化、主题演化分析。刘志辉等[12]提出作者关键词耦合网络(AKCA),这一模型可以发现作者之间的隐含的关系,并且随着作者发文量的增多,AKCA可通过关键词的耦合强度变化识别领域主题的变化。上述3种方法较为简单,能够适用于不同的领域,并且现阶段有多种成熟的分析工具(如Citespace、NEViewer)可用于文献计量分析。然而引文分析存在时滞性问题,同时还面临着链接数量庞大、耗费资源多的问题[13];共词分析对关键词的选择具有较强的敏感性[14],分析结果可能会因为词的选择带来较大差异;作者关键词耦合分析侧重于分析作者合作关系,通过作者的高频关键词间接获得研究主题。

基于概率主题模型的主题变迁方法在近些年获得越来越多的关注,李湘东等使用LDA模型和JS散度研究主题内容和主题强度随时间的变化[15]。倪丽萍等使用LDA模型识别不同时间片上的技术主题,并通过AP聚类算法对全局的主题进行聚类,形成领域的发展路径[16]。D. M. Blei等提出动态主题模型(Dynamic Topic Model)[17],认为主题随着时间不断变化,根据发表时间将文档划分到不同的时间片中,每一个时间片分别提取主题,再用KL散度计算不同时间片主题分布相似度。基于主题模型的方法能够解决引文分析中固有的时滞性问题,相比共词分析可以挖掘出深层次的语义信息,但是其揭示的只是统计概率层面的语义关系,并且主题个数需要提前指定,无法动态调整。

现阶段绝大多数主题变迁方法都是针对文献的内部特征,即文档内容本身进行分析,如何将学术文献中包含的外部特征(如作者、发表期刊等)融合进主题分析中,学者们针对不同的特征提出了相对应的算法模型。其中典型方法是Rosen-Zvi提出的AT模型(author-topic model)[18],通过在LDA模型基础上加上作者隐变量,AT可以获得作者-主题分布,但这一模型隐含的假设每个作者只有一个主题[19],不适应作者研究主题多样且不断变化的实际情况。D. M. Blei等提出RTM模型(Relational Topic Model)[20],通过将文档内容和文档之间的链接混合建模,实现更好的词汇预测和链接预测,但如果没有预测链接这一需求,RTM模型稍显复杂。利用上述两种主题模型可以更好地识别不同时间片中的主题,再根据KL距离或JS散度计算不同主题之间的相似度从而实现主题关联,经过关联过滤之后获得最终的主题变迁结果。

2.2 路径可视化

为了便于直观了解领域发展路径,主题变迁需要进行可视化展示。主题变迁和可视化关系密不可分,可视化是为了更好地进行主题变迁分析。现有的路径可视化方法众多,其中陈超美基于java开发出Citespace,具有引文分析、时序网络可视化等功能[21]。S. Havre等推出ThemeRiver模型,通过“河流”的宽度反映主题强度随着时间的变化[22]。微软亚洲研究院提出TextFlow方法,在文本分析中添加主题融合和分裂信息[23]。本文采用加权雅克比相似度进行主题关联,并借鉴了TextFlow方法的精髓,设计了和本文路径识别方法相适应的可视化方法。

综上所述,领域研究主题变迁包含多种因素,例如研究内容演变与学者变迁。因为专家学者是科学研究的主体,一些杰出学者往往引领着学科的发展,可以通过挖掘杰出学者公开发表的科技文献,发现科学领域的主要研究主题以及主题之间的联系[24]。关键词反映了学者的研究内容,一些关键词被某一学者在发表的文章中提及,可能表明这些关键词之间存在一定的关联关系[25];如果这些关键词被不同的作者提及,可能说明不同作者认可这些关键词之间的关联关系。相比共词分析,通过以作者为词共现单元构建关键词-作者矩阵,经过聚类得到的关键词簇具有更深层次关联关系。KMeans++适应海量文本聚类的场景,与谱聚类结合使用,相比主题数量固定的主题模型,能够实现动态调整类别个数,该方法具有一定的灵活性。采用相似度阈值法关联相邻时间段内的主题,并利用D3.js语言将主题变迁结果可视化展示,可以清晰地展示领域发展路径。

3 方法流程

图1是本方法的整体流程,主要包括数据来源与预处理、主题识别、主题关联和关联网络可视化4步。

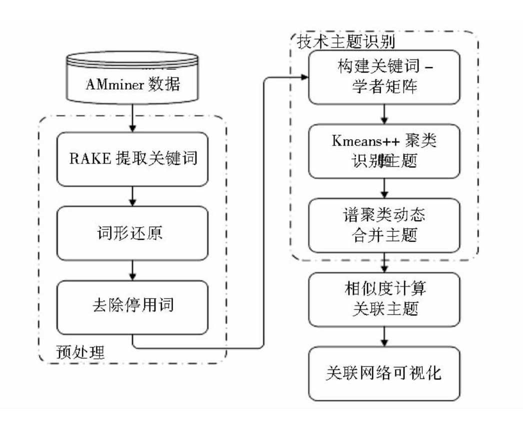


图 1　方法流程

3.1　数据来源与预处理

数据源选取 AMiner 科技数据库[26]，AMiner 通过数据挖掘和社会网络分析等方法，提供精准的领域专家推荐，并且该平台数据开源，可以通过 API 方便获取研究所需的专家学者及学者发表的论文数据。本文提出的一种基于关键词－专家的通用方法，通过 AMiner 可以确定领域相关的专家学者，通过 API 获取这些专家的所有发表的文献。数据预处理流程分为提取关键词、词形还原、去除停用词 3 个步骤。

（1）提取关键词。英文文献由于其语言特性，不需要进行分词处理。然而单个词汇表达的含义不具体，相对词汇而言，短语是更有意义的语义结构。本文使用 RAKE 算法[27]提取标题和摘要中长度为 2 个单词或 3 个单词的短语，作者提供的关键词直接添加进关键词列表中。

英文中的关键词通常由多个单词组成，关键词中不会有标点符号，也很少存在一些虚词，例如 an、this、but 等。RAKE 算法根据标点符号将文档分割成若干个子句，然后使用停用词将句子继续分成一些短语，并将这些短语作为潜在关键词。每个短语的分数是由构成短语的词累积而成：

$$score(w) = \frac{wordDegree(w)}{wordFrequency(w)} \qquad 式(1)$$

其中，*score* 为单词 *w* 的得分，*wordDegree* 表示为单词 w 的度（每当与一个单词共现时，度加 1），*wordFrequency* 表示为单词 *w* 在文档中出现的总次数。

（2）词形还原。英文单词具有多种形式，所以需要对关键词进行词形还原处理，即合并一些实际意义相同，但形式不同的单词。针对词形还原，本文采用 NLTK 中 Stemming 算法进行词形还原。

（3）去除停用词。词形还原之后，需要去除停用词：去除一些意思过于宽泛的关键词，例如“artifici intellig”“data mine”等；去除一些无意义的词汇，如“case studi”“data source”等。去除停用词之后，获得最终的关键词列表。

3.2　基于学者特征识别主题

为了识别不同时间段内的主题，首先需要将时间序列划分为若干个长度为 L 的时间片，根据发表时间，将文献划入到相应的时间片中。之前关键词聚类大多使用共词矩阵，本文尝试以学者为特征，即关键词为行向量，学者为列向量，使用 KMeans++ 算法[28]对单个时间片内的关键词进行聚类，得到的关键词簇可认为是时间片的研究主题。关键词可以利用向量空间模型（VSM）进行表示，其中每个学者为一个维度则每个关键词 t 映射成为：

$$v(t_j) = (a_1, t_jf_1; \cdots; a_i, t_jf_i; \cdots; a_n, t_jf_n) \qquad 式(2)$$

其中 $a_i(i=1,2,\cdots,n)$ 表示第 i 个作者，t_jf_i 表示为在第 i 个作者所有的文章中关键词 t_j 出现的频次。关键词－作者关联矩阵如下所示：

$$Matrix = \begin{array}{c} \\ t_1 \\ \\ t_j \\ \\ t_m \end{array} \begin{bmatrix} a_1 & & a_i & & a_n \\ t_1f_1 & \cdots & t_1f_i & \cdots & t_1f_n \\ \cdots & & \cdots & & \cdots \\ t_jf_1 & \cdots & t_jf_i & \cdots & t_jf_n \\ \cdots & & \cdots & & \cdots \\ t_mf_1 & \cdots & t_mf_j & \cdots & t_mf_n \end{bmatrix} \qquad 式(3)$$

3.2.1　利用 Kmeans++ 算法识别主题　KMeans++ 是基于传统 KMeans 算法的一种改进算法。KMeans 以距离作为划分类别的标准，通过 KMeans 聚出的类别具有同类别中相似度高、不同类别中相似度低的特点[29]。然而，KMeans 需要在聚类之前人工指定初始聚类中心，Kmeans++ 算法可以解决这一缺点，通过最大化初始聚类中心（质心）之间的距离，选出较优的聚类中心。KMeans++ 算法具有运算速度快、调参方便、易于理解结果的特点，适用于海量文本聚类的情形。

3.2.2　谱聚类合并相邻主题　使用基于图结构的谱聚类算法[30]，合并相似且同一时间片的节点，最终得到领域在时间片内的研究主题。不论是 Kmeans 算法还是 Kmeans++ 算法，始终面临着在聚类前需要事先给出初始的聚类个数这一问题，现实中很多时候即使是专家学者也不能确定该分多少类合适，但是可以确定一个大概的范围，即这个时间片内最多会有多少个主题。由此在初次聚类时需要专家确定主题数量范围，之后采用谱聚类算法，对 Kmeans++ 算法已经识别出的主题进行二次聚类，合并同一时间片内的相似主题得

到最终主题。

谱聚类算法一般流程：

(1)构造邻接矩阵。在谱聚类算法中，首先需要构建邻接矩阵 W(即主题相似度矩阵)[31]。采用加权雅克比相似度算法计算主题节点之间的相似度：以关键词的词频作为权重，逐一计算每个主题和所有主题之间的相似度，最终得到相似度矩阵。

(2)记 D 是度矩阵，计算 Laplacian 矩阵 $L = D\text{-}W$，并用式(4)将 Laplacian 矩阵标准化：

$$L_N = D^{-1/2} L D^{1/2} \qquad \text{式(4)}$$

(3)计算 L_N 的前 k 个最小特征值对应的特征向量 $v_1, v_2, \cdots, v_k$，令 $V = [v_1, v_2, \cdots, v_k]$，其中矩阵 V 的行数为主题节点个数，列数为 k。

(4)利用 KMeans 算法将矩阵 V 聚成 s 个类。

3.2.3 主题相关作者识别　为了探索领域中最重要的研究主题变化，本文选取领域的核心研究学者，并搜集这些学者发表的文章。如果某个关键词在某一学者的文章中多次出现，可以表明该学者对关键词所代表的研究主题做了大量研究，并间接说明该学者对该领域具有一定的影响力。

经过两次聚类，所有的关键词被分配到若干主题中。由于在关键词向量 $(a_1, t_j f_1; \cdots; a_i, t_{jfi}; \cdots; a_n, t_j f_n)$ 中 $t_j f_i$ 代表的是学者 a_i 发表的文章中提及该关键词 t_j 的次数，所以可以认为向量中 $t_j f_i$ 最大数值所对应的学者对关键词研究较多，并且该学者对关键词所代表的领域影响力较大。由此，通过计算每个主题的质心，根据质心向量中每一位数值大小，可以确定主题研究中影响力较大的学者。

3.3 基于相似度计算实现主题关联

经过两次聚类得到不同时间窗口的主题，然而这些主题都是相互独立的，为了进一步分析主题的变迁，需要对已经识别出的不同时间窗口的主题进行关联。主题变迁的实质是主题内容的改变，相邻时间片内的主题可以通过相似度计算得到关联关系。

此处使用加权雅克比相似度算法进行不同时间窗口主题之间相似度计算。记第 s 时间段的第 i 个主题为 T_i^s，第 $s+1$ 时间段的第 j 个主题为 T_j^{s+1}，$w(t, T_i^s)$ 为关键词 t 在时间段 s 中出现的频次：

$$w(T_i^s) = \sum\nolimits_{t \in T_i} w(t, T_i^s) \qquad \text{式(5)}$$

$$sim(T_i^s, T_j^{s+1}) = \frac{w(T_i^s \cap T_j^{s+1})}{w(T_i^s \cup T_j^{s+1})} \qquad \text{式(6)}$$

其中，

$$w(t, T_i^s \cap T_j^{s+1}) = min(w(t, T_i^s), w(t, T_j^{s+1})) \quad \text{式(7)}$$

$$w(t, T_i^s \cup T_j^{s+1}) = max(w(t, T_i^s), w(t, T_j^{s+1})) \qquad \text{式(8)}$$

本文得到的主题由一系列关键词组成，可以通过加权雅克比相似度计算不同主题之间的相似性，从而关联主题得到一系列主题对，由于这些主题对相似度高，可认为存在主题变迁关系。基于相似度计算会产生一些内容延续不明显的无效关联，所以需要对关联的主题进行过滤，从而突出核心的主题变迁。本文采用设定阈值法进行关联过滤：记第 s 时间段中的所有主题与 T_j^{s+1} 的相似度总和为 sum，ε 为相对阈值，如果 $sim(T_i^s, T_j^{s+1})/sum < \varepsilon$，则可认为关联无效。经过过滤的主题对有较强关联，能够表示主题变迁。

3.4 关联网络可视化设计

本文基于 D3. js 对关联结果进行可视化展示，便于科研人员和管理者直观地了解领域的发展动态。通过相似度计算，不同时间窗口的主题建立了关联，通过可视化处理，有助于理解与分析技术主题的发展。

主题变迁涉及 4 类信息：主题强度、主题内容、关联关系、主题相关学者。在可视化设计中，包括了点、线两种显示元素，其中点表示时间片上的主题，线表示主题之间的关联。为了显示更多的信息，本方法生成的领域发展路径可以通过 Web 查看，并添加了一定程度的交互：鼠标移动到主题点上，出现 Top5 的高频词和相关度最高的 5 名学者；鼠标按住节点可以实现拖动。

(1)主题强度。主题强度是研究主题的热度，本文识别出的主题是一组关键词的集合，因此主题强度采用主题中包含的关键词的词频总和进行度量。在可视化设计中，主题强度通过节点宽度 $node_width$ 表示：

$$node_width_i^s = \frac{w(T_i^s)}{\sum_{j \in s} w(T_j^s)} \times num_doc(s) \qquad \text{式(9)}$$

其中 $node_width_i^s$ 表示第 s 时间片上第 i 个主题的强度，$w(T_i^s)$ 表示该主题的总词频，$num_doc(s)$ 为第 s 时间片内论文的数量。

(2)主题内容。包括主题名称和主题内关键词，不同的关键词共同构成了主题的研究内容。本文采用主题中词频最高的关键词作为主题名称，并在弹框中显示 Top5 的高频词。

(3)关联关系。相邻时间片的主题之间相互连接，形成变迁关系。假定主题 T_i^s 和 T_j^{s+1} 能形成关联，由于相同的关键词在不同主题中的所占比例不同，连接线两端的粗细不同。

前向宽度为连接线在 $s+1$ 时间片上显示的宽度，

$$forward_width = node_width_j^{s+1} \times \frac{sim(T_i^s,\ T_j^{s+1})}{sim(T_j^{s+1})}$$

式(10)

其中 $sim(T_i^{s+1})$ 表示和主题 T_j^{s+1} 形成关联关系对应的相似度和。

后向宽度为连接线在 s 时间片上显示的宽度，

$$backward_width = node_width_i^s \times \frac{sim(T_i^s,\ T_j^{s+1})}{sim(T_i^s)}$$

式(11)

其中 $sim(T_i^s)$ 表示和主题 T_i^s 形成关联关系对应的相似度和。

(4)主题相关学者。弹框中显示在研究主题中影响力 Top5 的学者。

4　实验与结果分析

4.1　案例选择与数据处理

为了验证领域发展路径识别方法的有效性，本文以"artificial intelligence"为关键词，通过 API 从 AMiner 获取人工智能领域专家学者。人工智能发展时间较长，为了从复杂的领域发展中找出最核心的变化，本文选择领域 Top100 专家学者，之后同样通过 API 获取这些专家的全部论文数据。由于"人工智能"一词起源于 1956 年的达特茅斯会议[32]，本文限制论文的时间跨度从 1956 年到 2017 年，共计获得 25 614 篇文章。根据王丽雅发现的计算机领域论文的 5 年半衰期[33]，同时考虑田金萍总结的人工智能的 5 个发展阶段[34]，为了方便时间片划入对应发展阶段，本文根据发表时间将文献以 6 年为区间划分为 11 个时间片，其中 2016 –2017 单独组成一个时间片。各个时期的文献数量变化见图 2，1956 –1991 年间发文量平缓上升，1992 –2009 年间发文量呈指数式增长，反映了学者开始在这一领域发力，人工智能研究热度快速提高，其中 2016 –2017 时间片只包含两年时间，收录的文章数量较少。专家及论文数据可通过 API 获取(http://doc.aminer.org/en/latest/)。

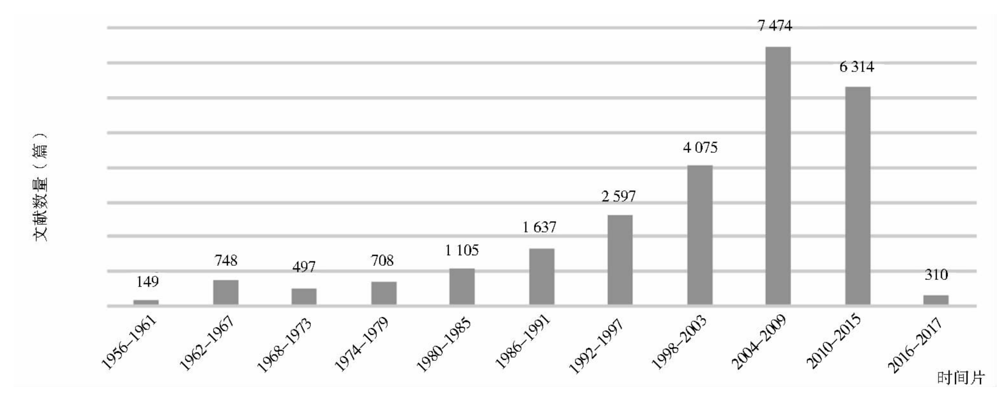


图 2　文献数量变化

4.2　主题识别与关联

获取文献之后首先进行预处理，通过 RAKE 算法提取关键词并进行词形还原，选取各时间片出现频率最高的前 1 000 个关键词，通过人工方式构建停用词表。为了看清领域的宏观主题变化，初次聚类 K 值可以设置小一些，此处定为 10，通过 Kmeans++ 算法求得各个时间片上的主题。通过谱聚类的方式合并类似的主题，谱聚类的 K 值同样设置为 10。初次聚类后如果某个时间片上 10 个主题相关度不高，经谱聚类后的该时间片主题数可以保持为 10；如果一些主题的相关度很高，经谱聚类可被合并成一个主题，从而实现了主题数量的动态调整。主题的识别结果见表 1。

11 个时间片共识别出 68 个主题，接下来使用基于加权雅克比相似度计算进行主题关联。经过尝试，当相似度阈值 ε 设置为 0.2 时显示效果最好，相似度大于阈值则判定主题对之间具有关联关系。

4.3　领域发展路径分析

经过可视化处理之后，人工智能领域发展路径见图 3。

主题变迁图中包含点和线两种元素，其中点表示某一时间片上具体的研究主题，线表示主题之间关联，线的粗细表示关联强弱。从左向右随着时间的推移，时间片上不断发生主题的产生、消亡、继承、融合和分裂。根据人工智能 5 个发展阶段的理论[34]，本文将 11 个时间片进行划分，其中[1956 –1961]为第一阶段，[1962 –1967]、[1968 –1973]为第二阶段，[1974 –1979]、[1980 –1985]、[1986 –1991]作为第三阶段，[1992 –1997]作为第四阶段，[1998 –2003]、[2004 –

周源，张超，唐杰，等.基于主题变迁的领域发展路径智能化识别——以人工智能为例[J].图书情报工作,2018,62(14):62－71.

表 1　主题识别结果

时间片	主题数量	主题群
1956－1961	5	inform retriev, associmemori, oper research, intellig system, inform theori
1962－1967	6	inform process, decistheori, mathemat, system analysi, kinet, automat control
1968－1973	6	semant, pattern recognit, heurist, formal languag, modal logic, set theori
1974－1979	6	semant network, naturlanguag, simul, decis support system, logic, control structur
1980－1985	7	inform retriev, expert system, data process, decis support system, knowledg represent, naturlanguag, pattern recognit
1986－1991	6	distribut process, expert system, neural network, computarchitectur, learn method, parallel process
1992－1997	7	expert system, computarchitectur, neural network, fuzzi logic, knowledgebas system, comput model, softwarengin
1998－2003	5	knowledg base system, fuzzi logic, mobil robot, neural network, semant web
2004－2009	7	predict model, neural network, adapt system, multiag system, inform retriev, semant web, pattern recognit
2010－2015	7	fuzzi logic, comput complex, multiag system, social media, neural network, predict, bioinformat
2016－2017	6	knowledgmanag, bayesian network, object function, graphic model, real world, neural network

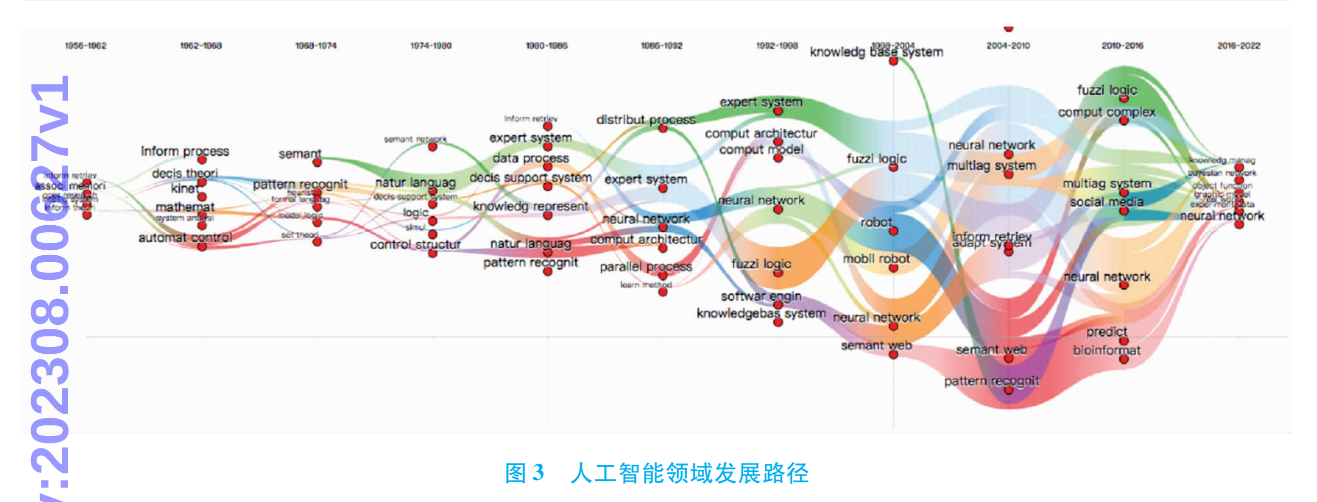


图 3　人工智能领域发展路径

2009]、[2010－2015]、[2016－2017]作为第五阶段。由于[2016－2017]时间片只包含两年时间，收录的文章数量较少，而在可视化设计逻辑中节点宽度和时间片文章的数量成正相关，主题变迁图在最后一个时间节点上呈现“收敛”的效果。人工智能领域发展路径图可以通过网页访问(网址：http://118.24.155.51:8080/trend_ai/)。网页中节点可以拖动，鼠标停留在节点上会显示节点的核心的 5 个主题关键词和主要的 5 个专家学者。

由于第三阶段在人工智能的发展历程中起到承前启后的作用，限于篇幅，本文选取第三阶段(1974－1991)进行深入的主题内容的分析。第三阶段的放大图见图 4。

从图 4 可以直观地看出第三阶段的研究热点，包括专家系统(expert system)、神经网络(neural network)、知识表示(knowledge represent)、自然语言处理(naturlanguag)、并行处理(parallel process)等。

4.3.1　主题强度变化、关联分析　专家系统在第三阶段中出现了两次且线条较粗，其中 T2 中的专家系统和

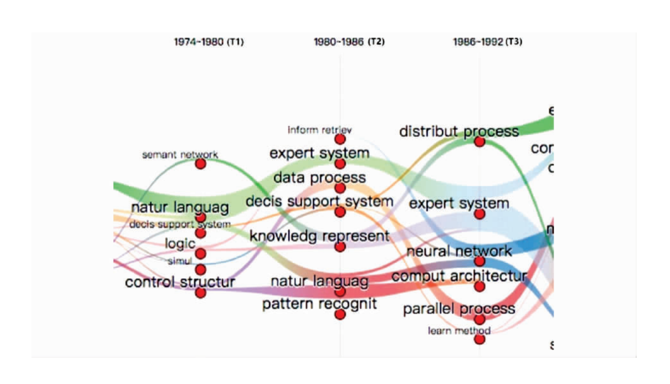


图 4　第三阶段主题

T1 中的自然语言处理强相关，T3 中的专家系统和 T2 中的知识表示强相关，这说明专家系统融合了多项之前的研究结果，成为这段时间的核心研究主题。现实中，20 世纪 60 年代到 70 年代末，由斯坦福大学开发出的 DENDRAL 质谱分析系统标志着专家系统的出现，人工智能的研究由此进入新领域。神经网络在 T3 时间片上出现，其与 T2 中决策支持系统、自然语言处理和信息抽取相关联，说明神经网络的研究得益于这三类学科的发展。知识表示在 T2 时间片出现，20 世纪

80 年代,日本启动能够进行大规模并行处理的第五代计算机研制计划,有关知识工程的研究开始进入繁盛期。T2 时间片上 pattern recognit(模式识别)和前后时间片上的研究主题关联较弱,在图中表示为一个孤立的点。

4.3.2 主题内容变迁分析 通过两次聚类的方式获得了每个时间窗口内研究的核心主题和主题相关关键词,选取每个主题出现频次 Top5 的关键词表示主题的详细内容,对主题进行深入分析。第三阶段所有主题如表 2 所示:

表 2 第三阶段所有主题

时间片	主题名称	关键词
T1	control structure	control structur, theorem, product system, dynam program, human expert
	Logic	logic, fuzzi logic, knowledg represent, membership function, infer rule
	decision support systems	decis support system, computsimul, linear program, optim, comput complex
	simulation	simul, distributcomput, game theori, manag, distribut system
	semant network	semant network, data structur, data manag, databasmanag, data model
	nature language	naturlanguag, semant, pattern recognit, comput linguist, inform process
T2	expert system	expert system, rule base, spectrum, knowledgengin, knowledgacquisit
	data processing	data process, computarchitectur, distributcomput, parallel process, comput network
	decision support systems	decis support system, manag, oper research, decis support, computapplic
	knowledg represent	knowledg represent, conceptu model, fuzzi logic, knowledg base system, probabl
	nature language	naturlanguag, reason, mathemat model, inform process, problem solv
	inform retrieve	inform retriev, manag system, user interfac, knowledgmanag, inform analysi
	pattern recognition	pattern recognit, product system, imag process, dynam program, heurist search
T3	distribution process	distribut process, softwarengin, oper system, decis support system, architectur design
	expert system	expert system, knowledg represent, knowledg base system, logic program
	neural network	neural network, inform retriev, intellig system, bioinformat, biomed research
	computer architecture	computarchitectur, cognit, robot, mobil robot, network
	learning method	learn method, bayesian network, converg, genet, search space
	parallel processing	parallel process, data structur, comput network, schedul, real time system

模型能够识别出的每个时间片的研究主题。从表 2 可以看出:每个主题的方向较为明确,且其中的关键词能够表明主题的研究内容。以 T3 中的神经网络为例,其中 neural network(神经网络)、inform retriev(信息抽取)、intellig system(智能系统)、bioinformat(生物信息)、biomed research(生物医学研究)这 5 个词出现频次最高,说明最初的神经网络和生物技术研究关系比较紧密,这反映了一个事实:神经网络起源于人类对于大脑神经元的研究。智能系统说明了神经网络最初的应用场景。

从表 2 中可以看出随着时间的推进研究主题内容不断变化。以专家系统为例,T2 中专家系统中包含 rule base(基于规则)、spectrum(质谱)、knowledgengine(知识引擎)、knowledgacquisit(知识获取),表明由质谱分析发展而来的一些类似系统共同组成了新一代专家系统,此时的专家系统基于规则,集成了知识。从 T2 至 T3 的发展过程中,T2 中的专家系统和知识表示发生融合形成新一代专家系统,包含 knowledg represent(知识表示)、knowledg base system(基于知识的系统)、logic program(逻辑程序)、knowledg engine(知识引擎),此时的专家系统更加注重知识。从两个阶段的专家系统的细微变化可以看出,专家系统由开始的规则导向逐渐演化成为知识导向。

4.3.3 主题相关学者变迁分析 模型基于学者对关键词进行聚类,每一个类的质心指向该主题的核心学者,由此可以看出在研究主题融合和分裂的过程中学者的变化,学者的变化可以从侧面反映主题内部研究方向的变化,图 5 所示的弹框中包含了主题的关键信息和核心学者。

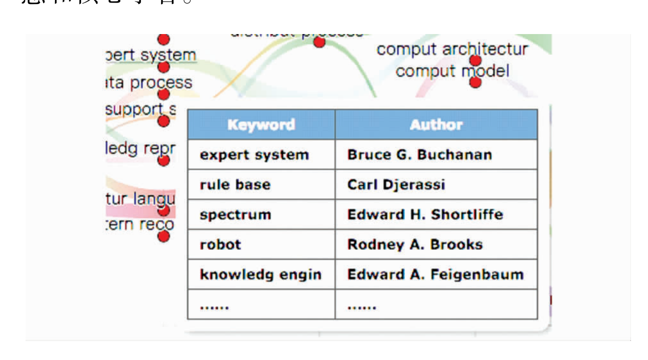


图 5 主题节点内容

选取专家系统有关的学者进行进一步分析,专家系统作为一个单独的主题在 1980 – 1986 时间片首次出现,在 1992 – 1998 时间片最后一次出现,共计在 3

个时间阶段出现。如表 3 所示：

表 3　专家系统相关学者

时间片	核心学者
1980－1985	B. G. Buchanan, C. Djerassi, E. H. Shortliffe, R. A. Brooks, E. A. Feigenbaum
1986－1991	H. Prade, D. Dubois, P. R. Cohen, Y. Wilks, S. C. Shapiro
1992－1998	N. R. Jennings, R. A. Brooks, E. Horvitz, A. K. Mackworth, T. M. Mitchell

1980－1985 时间片的 5 名专家学者中，B. G. Buchanan 是匹兹堡大学计算机科学系教授，他与 E. A. Feigenbaum 等科学家共同研发了第一代专家系统——DENDRAL，这是一款应用于质谱分析（spectrum）的专家系统[35]。C. Djerassi 在化学研究中大量应用质谱分析，扩展了专家系统的应用场景[36]。E. H. Shortliffe 的研究和医疗相关，他在本时间片内发表了大量和医疗咨询系统设计与专家系统性能优化相关的文章[37]。由此，从专家的研究方向可以看出该时间片内专家系统高度专业化，以解决某一领域的特定问题为目标。

1986－1991 时间片中，H. Prade、D. Dubois 的研究和模糊理论相关，H. Prade 发表了基于可能性理论提升原有基于规则的专家系统的文章[38]。P. R. Cohen、Y. Wilks、S. C. Shapiro 三人的研究聚焦于信息抽取、自然语言理解，P. R. Cohen 改进 GRANT 专家系统存在的性能下降问题[39]。第二代专家系统引入了信息抽取、模糊理论和不确定推理技术，通用性更强，一定程度上解决了第一代专家系统过度专业化的弊端。

1992－1998 时间片中，N. R. Jennings 的研究和多智能体系统（multi agent system）相关，该系统用于解决单个专家系统难以解决的问题[40]。E. Horvitz 在此时间片内进行决策理论、可能性模型相关研究[41]。T. M. Mitchell 是人工智能领域的核心学者，他在本时间片中发表了从互联网上提取知识相关的文章[42]。这一阶段核心专家的研究方向较为分散，可以看出传统的专家系统研究日渐式微，其研究的内容逐渐转向大规模知识获取、知识表示以及多专家系统协作。

4.4　领域发展路径识别有效性验证

验证方法的有效性，是主题变迁研究中普遍存在的一个难题。对于主题内容变迁和主题强度变迁，现阶段并没有通用的标准能够对结果的有效性进行评估；对于不同的主题变迁模型，也缺乏有效的定量比较方法[43]。除了定量指标之外，笔者更加关心的是识别出的主题变迁是否能够真实反映领域发展的实际情况，因此召开了专家咨询会议对结果进行评估，与会专家包括 2 位副高级职称计算机领域专家、1 位博士研究生和 2 位硕士研究生。

经过评估，专家认为识别出的人工智能领域发展路径总体结果准确，不同时间片间主题和主题之间的关联较为合理，可视化展示较为清晰，对于研判人工智能领域技术发展具有一定的参考价值。但同时存在着一些问题：①一些主题出现的时间节点有偏差；②在近期的研究中，一些主题被遗漏，例如语音识别，计算机视觉；③主题之间具有层次关系，但是在图中所有的主题都是并列排布。

经过分析笔者认为：主题出现时间有偏差，可能是由于知识扩散需要一定的时间，导致专家所做的评判和与基于数据挖掘而做的评判存在一定的差异。对于近期一些主题被遗漏的情况，笔者在原始数据中对这些主题进行手动检索，发现相关的文章数量偏少，可能是由于这些主题学界研究较少，但在工业界研究较多，针对这一情况需要引入其他数据（例如专利）对结果进行修正。针对主题之间的层次关系，笔者在未来的工作中将尝试引入 Wiki 百科树形结构知识，在聚类时考虑关键词的粒度信息。经过分析笔者发现，结果中有不符合专家认知的情况，主要因为数据源单一造成了一定程度的偏差，因此在未来的研究中我们将会对数据源进行扩展，利用多源数据优化领域发展路径的结果。

5　结论

本文采用两层聚类与相似度计算相结合的方法，探索领域主题变迁，利用 D3. js 语言对主题变迁结果进行可视化展示，并以人工智能领域为例对该方法的有效性进行验证。

本研究的意义在于：①提出基于主题变迁的领域发展路径识别方法，在常用的主题识别方法基础上引入学者信息，从多视角更加全面地分析技术主题变迁，为领域发展路径相关研究提供新的方向；②本文所提出的方法可以用于分析领域技术发展规律，帮助科研人员快速定位领域发展热点和重要学者，辅助科技决策；③相比传统方法，本文提出的方法更加方便快捷，通过输入领域名称，可全自动地输出领域发展路径图。该方法将产品化上线，帮助更多的人直观地了解科学发展动态，接触领域发展前沿。

然而，本文把时间片大小设置为不可变，未考虑领域发展的速度，例如由于计算能力的限制，人工智能前期发展缓慢，此时时间跨度应该设大一些；近些年来随

着云计算、大数据的广泛应用,人工智能发展迅速,应该以细粒度划分时间窗口。因此在接下来的研究中,笔者将根据领域的实际情况,灵活设定不同领域的时间窗口大小,从而提高主题变迁分析的准确性与合理性。本文预处理过程中去停用词还停留在人工处理的阶段,需要找到针对不同领域的通用设置停用词方案。另外还需要引入多源数据,解决单一数据造成的局限性。

参考文献:

[1] RADZIEMSKI L J. From LASER to LIBS, the path of technology development[J]. Spectrochimica Acta part B: atomic spectroscopy, 2002, 57(7): 1109 – 1113.

[2] 王纬. 德尔菲调查法与技术路线图结合的技术预测研究——以太原市“十二五”技术发展预测为例[J]. 中国科技论坛, 2011(4):103 – 107.

[3] 徐骥, 张卫国, 罗军. 基于技术路线图分析法和 AHP 的企业技术发展路径规划[J]. 科学学与科学技术管理, 2010, 31(11):13 – 18.

[4] 潘颖. 基于专利引证强度的关键技术发展路径研究[J]. 情报理论与实践, 2014, 37(12):71 – 75.

[5] 许冠南, 谢梦娇, 潘美娟,等. 3D 打印产业技术的演变与预测研究——基于专利主路径分析[J]. 北京邮电大学学报(社会科学版), 2016, 18(4):77 – 85.

[6] 韩毅, 金碧辉. 基于连通性的引文网络结构分析新视角:主路径分析[J]. 科学学研究, 2012, 30(11):1634 – 1640.

[7] 秦晓慧, 乐小虬. 基于 LDA 主题关联过滤的领域主题演化研究[J]. 现代图书情报技术, 2015, 31(3):18 – 25.

[8] 赵迎光, 洪娜, 安新颖. 主题模型在主题演化方法中的应用研究进展[J]. 现代图书情报技术, 2014, 30(10):63 – 69.

[9] 洪宇, 张宇, 刘挺,等. 话题检测与跟踪的评测及研究综述[J]. 中文信息学报, 2007, 21(6):71 – 87.

[10] 崔雷, 王孝宁. 学科主题演变的深度挖掘分析——以普通外科学为例[J]. 医学信息学杂志, 2009, 30(8):5 – 10.

[11] 唐果媛, 张薇. 基于共词分析法的学科主题演化研究进展与分析[J]. 图书情报工作, 2015, 59(5):128 – 136.

[12] 刘志辉, 张志强. 作者关键词耦合分析方法及实证研究[J]. 情报学报, 2010, 29(2):268 – 275.

[13] 伍若梅, 孔悦凡. 共词分析与共引分析方法的比较研究[J]. 情报资料工作, 2010(1):26 – 29.

[14] 李纲, 巴志超. 共词分析过程中的若干问题研究[J]. 中国图书馆学报, 2017, 43(4):93 – 113.

[15] 李湘东, 张娇, 袁满. 基于 LDA 模型的科技期刊主题演化研究[J]. 情报杂志, 2014(7):115 – 121.

[16] 倪丽萍, 刘小军, 马驰宇. 基于 LDA 模型和 AP 聚类的主题演化分析[J]. 计算机技术与发展, 2016, 26(12):6 – 11.

[17] BLEI D, LAFFERTY J D. Dynamic topic models[C]//Proceedings of the 23rd international conference on machine learning. Pittsburgh:ACM, 2006: 113 – 120.

[18] ROSEN-ZVI M, GRIFFITHS T, STEYVERS M, et al. The author-topic model for authors and documents[C]//Proceedings of the 20th conference on uncertainty in artificial intelligence. Arlington: AUAI Press, 2004: 487 – 494.

[19] 史庆伟, 乔晓东, 徐硕,等. 作者主题演化模型及其在研究兴趣演化分析中的应用[J]. 情报学报, 2013, 32(9):912 – 919.

[20] CHANG J, BLEI D. Relational topic models for document networks [C]//Proceedings of the artificial intelligence and statistics. Clearwater Beach:JMLR. org,2009: 81 – 88.

[21] 李杰,陈超美. Citespace:科技文本挖掘及可视化[M]. 北京:首都经济贸易大学出版社, 2016.

[22] HAVRE S, HETZLER E, WHITNEY P, et al. ThemeRiver: visualizing thematic changes in large document collections[J]. IEEE transactions on visualization & computer graphics, 2002, 8(1):9 – 20.

[23] CUI W, LIU S, TAN L, et al. TextFlow: towards better understanding of evolving topics in text[J]. IEEE transactions on visualization & computer graphics, 2011, 17(12):2412 – 2421.

[24] 刘萍, 郭月培, 郭怡婷. 利用作者关键词网络探测作者相似性[J]. 现代图书情报技术, 2013(12):62 – 69.

[25] LI H, ABE N. Word clustering and disambiguation based on co-occurrence data[C]// International conference on computational linguistics. Montreal: Association for Computational Linguistics, 1998:749 – 755.

[26] TANG J. AMiner: mining deep knowledge from big scholar data [C]// International conference companion on world wide Web. San Francisco: International World Wide Web Conferences Steering Committee, 2016:373 – 373.

[27] ROSE S,ENGEL D, CRAMER N, et al. Automatic Keyword Extraction from Individual Documents[M]//BERRY M W, KOGAN J. Text mining: applications and theory. Chichester:John Wiley & Sons,Ltd, 2010:1 – 20.

[28] ARTHUR D, VASSILVITSKII S. k-means++ : The advantages of careful seeding[C]//Proceedings of the eighteenth annual ACM-SIAM symposium on discrete algorithms. Philadelphia:Society for Industrial and Applied Mathematics, 2007: 1027 – 1035.

[29] 周爱武, 于亚飞. K-Means 聚类算法的研究[J]. 计算机技术与发展, 2011, 21(2):62 – 65.

[30] NG A Y, JORDAN M I, WEISS Y. On spectral clustering: analysis and an algorithm[C]//Advances in neural information processing systems. Vancouver:NIPS Foundation,2002: 849 – 856.

[31] 刘志伟. 谱聚类中的相似度矩阵研究[J]. 现代计算机, 2010(15):67 – 69.

[32] 吕泽宇. 人工智能的历史、现状与未来[J]. 信息与电脑, 2016(13):166 – 167.

[33] 王丽雅. 基于 CNKI 的计算机科学学科半衰期分析[J]. 图书与情报, 2015(1):100 – 105.

[34] 田金萍. 人工智能发展综述[J]. 科技广场, 2007(1):230 – 232.

[35] LINDSAY R K, BUCHANAN B G, FEIGENBAUM E A, et al. DENDRAL: a case study of the first expert system for scientific hypothesis formation [J]. Artificial intelligence, 1993, 61(2):209－261.

[36] SAMPLE S, DJERASSI C. Mass spectrometry in structural and stereochemical problems[J]. Journal of the American Chemical Society, 1966, 88(9):1937－1943.

[37] SUWA M, SCOTT A C, SHORTLIFFE E H. An approach to verifying completeness and consistency in a rule-based expert system [J]. AI magazine, 1982, 3(4): 16.

[38] FARRENY H, PRADE H, WYSS E. Approximate reasoning in a rule-based expert system using posibility theory: a case study[M]. Paris:Laboratoire des Langages et systèmesinformatiques,1985:407－414.

[39] KJELDSEN R, COHEN P R. Evolution and performance of the grant system[J]. IEEE Expert-intelligent systems and their applications, 1987, 2(2): 73－79.

[40] JENNINGS N R. On agent-based software engineering[J]. Artificial intelligence, 2000, 117(2): 277－296.

[41] PAEK T, HORVITZ E. Uncertainty, Utility, and misunderstanding: a decision-theoretic perspective on grounding in conversational systems[J]. Proceedings of the aaai fall symposium on psychological models of communication in collaborative systems. Cape Cod: AAAI Press,1999:85－92.

[42] CRAVEN M, DIPASQUO D, FREITAG D, et al. Learning to extract symbolic knowledge from the World Wide Web[J]. Coastal management, 1998, 31(2):121－126.

[43] ZHOU H, YU H, HU R. Topic evolution based on the probabilistic topic model: a review[J]. Frontiers of computer science, 2017, 11(5): 786－802.

作者贡献说明：

周源：负责框架设计、论文修改及撰写指导；

张超：负责文献调研与论文撰写；

唐杰：负责框架设计；

刘宇飞：负责观点提炼、论文修改；

张宇韬：负责程序编写与相关实验、论文修改。

Intelligent Identification of Field Development Trajectory Based on Topic Evolution: A Case Study of Artificial Intelligence

Zhou Yuan[1] Zhang Chao[2] Tang Jie[3] Liu Yufei[4] Zhang Yutao[3]

[1] School of Public Policy and Management, Tsinghua University,Beijing 100084

[2] School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074

[3] Department of Computer Science and Technology, Tsinghua University,Beijing 100084

[4] The CAE Center for Strategic Studies, Chinese Academy of Engineering, Beijing 100088

Abstract: [**Purpose/significance**] Identifying the trajectory of development is of great importance to scientific and technological innovations. However, existing methods such as expert interviews and citation analysis cannot meet the current situation of the explosive growth of literature. In response to this problem, this paper proposes a new identification method of filed development trajectory. [**Method/process**] This method identifies the research topics and related scholars by using Kmeans++ and spectral clustering algorithms with the keyword-scholar matrix, calculates the correlation between different topics, and finally visualizes the trajectory of developmen. [**Result/conclusion**] Through the empirical analysis of the field of artificial intelligence, the results show that the method can effectively reflect the evolution of the topic of field research, help researchers quickly locate popular research topics and focuses, and enrich the research methods related to the trajectory of field development .

Keywords: field development trajectory　topic evolution　KMeans++　spectral clustering　artificial intelligence